SCIENCE CHINA
Information Sciences

• LETTER •

February 2021, Vol. 64 129204:1–129204:3
https://doi.org/10.1007/s11432-018-9613-2

Improved sparse representation based on local preserving projection for the fault diagnosis of multivariable system

Qiu TANG[1,2*], Benqi LI[3], Yi CHAI[1,2], Jianfeng QU[1,2] & Hao REN[1,2]

[1]*College of Automation, Chongqing University, Chongqing* 400030*, China;*
[2]*Key Laboratory of Complex System Safety and Control, Ministry of Education, Chongqing* 400030*, China;*
[3]*Xichang Satellite Launch Center, Xichang* 615000*, China*

Received 29 June 2018/Accepted 13 September 2018/Published online 11 June 2020

Citation Tang Q, Li B Q, Chai Y, et al. Improved sparse representation based on local preserving projection for the fault diagnosis of multivariable system. Sci China Inf Sci, 2021, 64(2): 129204, https://doi.org/10.1007/s11432-018-9613-2

Dear editor,

To satisfy the increasing requirements for safety and quality in industrial processes, process monitoring has been actively investigated in the past decade. The most critical aspects of this approach are the detection of faults in real-time and the diagnosis of fault types. Redundancy and coupling among these variables make it difficult to identify existing correlations between faults and variables, which hinders the quick detection of faults. Furthermore, the large amount of available monitoring data often obscures information about abnormalities and faults. Effective dimension reduction and feature extraction are imperative to address this challenge [1, 2].

Traditional multivariate statistical methods such as principal component analysis (PCA), partial least squares (PLS), and independent component analysis (ICA), have been widely used in process monitoring. These dimension reduction methods commonly extract features based on the assumption that the data follow a particular distribution. As such, noise and measurement error in the practical process usually result that data do not strictly follow the expected distribution. Correspondingly, the accuracy of fault detection and diagnosis is significantly degraded. To solve this problem, some modifications and improvements are proposed, including kernel techniques and recursive methods that lead to kernel PCA, kernel PLS, and recursive PCA. Sparse techniques perform well in facilitating more interpretative results and excluding redundancy [3].

However, the aforementioned methods only preserve the global structure of the dataset but ignore the local neighborhood relations. Huang et al. [4] combined the mean of the dataset with local preserving projection (GLPP) to simultaneously preserve the local and global structures. The results for 6 datasets validated the effectiveness of the method. Zhan et al. [5] proposed a novel ensemble global-local preserving projection (GLPP) method, which introduced the Bayesian inference and weighted sum strategies to combine the separated GLPP models. The validity and effectiveness were verified by applying the approach to the Tennessee Eastman (TE) process.

This investigation focuses on a quick and accurate method to detect faults and identify fault types. Because the sparse approach can effectively exclude redundancy and allow more interpretative features to be obtained, the integration of sparse techniques and local preserving projection can preserve both local and global structures of the sample data.

As stated in [3], sparse representation was initially proposed by Wright et al. and is currently widely used in computer vision and pattern recognition. This technique not only affords high-fidelity representation of the observed signal in a more concise form, but also facilitates the mining of latent information in the data samples. In this study, the dataset is denoted as a matrix X ($X \in \mathbb{R}^{m\times n}$), in which m and n are the number of variables and samples, respectively. It consists of K classes.

For the data sample x, the sample can be represented by the learned dictionary and sparse coefficient as follows:

$$x = Ay, \tag{1}$$

where A is an over-complete dictionary of the samples. This means that all the samples in X can be represented by A with an appropriate sparse coefficient. There are many methods for learning dictionaries such as KSVD (K-singular value decomposition), MOD (method of optimal directions), and sparseNet. In this study, we utilize the locality preserving projection (LPP) method to train the dictionary. The sparse coefficient y is learned by relaxing the $L0$ norm to $L1$ norm optimization because solving the $L0$ norm is an NP-hard problem. The object function is

$$J(\hat{y}) = \arg\min_{y}\{||x - Ay||_2^2 + \lambda||y||_1\}, \tag{2}$$

* Corresponding author (email: Ltangqiu@gmail.com)

where $||\cdot||_2$ is the $L2$ norm, $||\cdot||_1$ is the $L1$ norm, and λ is the sparsity penalty to balance the significance of sparsity and accuracy of the reconstructed signal. Then, the raw data can be represented as

$$x = A\hat{y} + e, \tag{3}$$

where e is the reconstruction error, and $\hat{y}$ is the sparse coefficient.

LPP is a representative manifold-based dimension reduction method proposed by He [6]. The main advantage of this approach over traditional dimension reduction methods, such as PCA and linear discriminant analysis (LDA), is that it does not impose any sample distribution assumptions and can handle nonlinear cases. The LPP approach seeks to project the samples into another space with a lower dimension, in which the manifold structure of the samples is preserved. The objective function is described as follows:

$$Y = D^{\mathrm{T}}X, \tag{4}$$

where $D = [d_1, d_2, \dots, d_l] \in \mathbb{R}^{m\times l}$ is the preserving projection matrix, which is solved by the objective function defined as

$$J(d) = \arg\min_d \left\{ \sum_{i,j} (y_i - y_j)^2 w_{ij} \right\}, \tag{5}$$

in which y_i is the projection of x_i by $y_i = D^{\mathrm{T}}x_i$, w_{ij} is the i,jth entry of the adjacency weight matrix W, so that the projections can preserve the local structure of the dataset. w_{ij} is defined as follows:

$$w_{ij} = \begin{cases} \mathrm{e}^{-\frac{||x_i - x_j||^2}{\sigma}}, \text{if } x_j \in \Omega_k(x_i) \text{ or } x_i \in \Omega_k(x_j), \\ 0, \text{ otherwise}, \end{cases} \tag{6}$$

where σ is a constant parameter, and $\Omega_k(x)$ is the neighborhood of x that is defined by the k nearest neighbors. Eq. (5) can be rewritten as

$$\begin{aligned} J(d) &= \arg\min_d \left\{ \sum_{i,j} (y_i - y_j)^2 w_{ij} \right\} \\ &= \arg\min_d \left\{ 2\left(\sum_i y_i H_{ii} y_i^{\mathrm{T}} - \sum_{i,j} y_i w_{ij} y_i^{\mathrm{T}} \right) \right\} \\ &= \arg\min_d \left\{ 2\left(\sum_i d^{\mathrm{T}} x_i H_{ii} x_i^{\mathrm{T}} d - \sum_{ij} d^{\mathrm{T}} x_i w_{ij} x_j^{\mathrm{T}} d \right) \right\} \\ &= \arg\min_d \{2(d^{\mathrm{T}} X(H - W)X^{\mathrm{T}} d)\} \\ &= \arg\min_d \{2(d^{\mathrm{T}} XMX^{\mathrm{T}} d)\}, \end{aligned} \tag{7}$$

where H is a diagonal matrix consisting of $H_{ii} = \sum_j w_{ij}$, and $M = H - W$ is the Laplacian matrix. Furthermore, a constraint is imposed on LPP for normalization. The objective function of LPP is then defined as

$$\begin{aligned} &J(d) = \arg\min_d \{d^{\mathrm{T}} XMX^{\mathrm{T}} d)\}, \\ &\text{s.t.} \quad d^{\mathrm{T}} XDX^{\mathrm{T}} d = 1. \end{aligned} \tag{8}$$

The optimization problem is then transformed to a generalized eigenvalue problem as follows:

$$d^{\mathrm{T}} XMX^{\mathrm{T}} d = \lambda XDX^{\mathrm{T}} d. \tag{9}$$

The optimal LPP projection matrix D consisting of the eigenvectors $d_1, d_2, \dots, d_l$ is sorted by the corresponding eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_l$. Sparse LPP (SLPP) based process monitoring and fault diagnosis is described as follows. Let $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m\times n}$ denote a normalized training dataset. The SLPP method involves determining the optimal transform matrix $D_i = [d_1, d_2, \dots, d_l]$, which is regarded as the learned dictionary. Because there are K classes, the entire structured dictionary is $D_{\text{all}} = [D_1, D_2, \dots, D_K]$, which reconstructs the samples with sparse representation. This method integrates sparse representation and LPP, thereby preserving both the local and global structures of the dataset. The monitoring model is then constructed as follows:

$$y = D^{\mathrm{T}}X, \quad x = D^{+\mathrm{T}}y + e. \tag{10}$$

Reconstructing the samples with $d_1, d_2, \dots, d_l$, respectively, the class with the least reconstruction error e is the fault type of the samples. In this study, we introduce T^2 and SPE to detect the appearance of faults. The T^2 measures the variations in the model subspace, and SPE measures the variations in the residual subspace. They are defined as

$$\begin{aligned} &\text{SPE} = ||e||^2 = ||x - D^{+\mathrm{T}}y||^2, \\ &T^2 = y^{\mathrm{T}} S^{-1} y. \end{aligned} \tag{11}$$

The proposed method for fault diagnosis is applied to the Tennessee Eastman process, which is an industrial process with 12 manipulated variables, 19 composition measurements, and 22 continuous measurements. Normal datasets are generated under typical operating conditions, and 21 fault datasets are generated under the 21 faulty operating conditions. Sparse discriminant analysis [7] (SDA) and sparse exponential discriminant analysis [8] (SEDA) are introduced to facilitate a comparison with the improved sparse local preserving projection (ISLPP). The evaluation indexes of the diagnosis performance in this study are defined by the classification accuracy as follows:

$$\begin{aligned} &\text{Acc}(i) = \frac{n_i^{\mathrm{T}}}{n_i} \times 100\%, \\ &\text{anvA} = \frac{1}{K} \times \sum_{i=1}^{K} \text{Acc}(i), \end{aligned} \tag{12}$$

where $\text{Acc}(i)$ is the classification accuracy of the ith class, n_i^{T} and n_i are the number of correctly classified samples and the total number of samples in the ith class, respectively, anvA is the average classification accuracy of all types of faults, and K is the number of fault types.

Experiments based on the ISLPP method are performed with 15 types of faults, which include sticking, step, random, slow draft and unknown faults. The classification accuracy is shown in Table 1. It is seen that the accuracy of the ISLPP approach is higher than that of the SDA and SEDA methods. Moreover, most of the results show significant improvement. For example, the diagnosis accuracy of SDA for faults #4 and #19 are only 13.02% and 7.5%, respectively. Using the proposed method, it represents an enhancement of 72.7083% and 67.0833%, respectively. The local preserving projection maintains the local structure of the dataset and the sparsity constraint precludes coupling among variables and preserves the global structure of the data.

Table 1 Classification accuracy (%) of 15 faults in TE process with different methods

	ISLPP	**SDA**	**SEDA**
Fault #1	82.71	81.67	81.46
Fault #2	82.08	79.37	80.42
Fault #3	17.71	8.54	11.46
Fault #4	72.71	13.02	20.21
Fault #5	82.08	57.6	35.63
Fault #6	83.33	75.42	77.19
Fault #7	83.33	80.10	76.98
Fault #8	43.96	29.48	29.27
Fault #12	37.92	19.27	20.21
Fault #13	34.79	27.81	25.62
Fault #15	17.71	15.42	13.44
Fault #17	51.67	44.79	46.98
Fault #18	70.42	38.33	62.40
Fault #19	67.08	7.5	21.67
Fault #20	57.29	40.21	34.79

In this investigation, ISLPP is proposed by integrating the sparsity constraint with the local preserving projection method, and dictionary learning is invoked to solve the project matrix. The simulation is implemented in a MATLAB 2016a environment using a personal computer with an i5-6500 core CPU operating at 3.20 GHz, with 8 GB of RAM. The total time to train the project matrix and identify fault types was 267 s.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61773080, 61673076, 61633005).

References

1 Ge Z Q. Review on data-driven modeling and monitoring for plant-wide industrial processes. Chemometr Intell Laboratory Syst, 2017, 171: 16–25

2 Ge Z Q, Song Z, Ding S X, et al. Data mining and analytics in the process industry: the role of machine learning. IEEE Access, 2017, 5: 20590–20616

3 Tang Q, Chai Y, Qu J, et al. Fisher discriminative sparse representation based on DBN for fault diagnosis of complex system. Appl Sci, 2018, 8: 795

4 Huang S, Elgammal A, Huangfu L, et al. Globality-locality preserving projections for biometric data dimensionality reduction. In: Proceedings of Computer Vision and Pattern Recognition Workshops. New York: IEEE, 2014. 15–20

5 Zhan C J, Li S H, Yang Y P. Enhanced fault detection based on ensemble global-local preserving projections with quantitative global-local structure analysis. Ind Eng Chem Res, 2017, 56: 10743–10755

6 He X. Locality preserving projections. Adv Neural Inf Process Syst, 2003, 16: 186–197

7 Clemmensen L, Hastie T, Witten D, et al. Sparse discriminant analysis. Technometrics, 2011, 53: 406–413

8 Yu W, Zhao C. Sparse exponential discriminant analysis and its application to fault diagnosis. IEEE Trans Ind Electron, 2018, 65: 5931–5940